Efficient Optimization Algorithms for Multi-User Beamforming With Superposition Coding

Efficient Optimization Algorithms for Multi-User Beamforming With Superposition Coding

Xiaoyan Shi, John S. Thompson, *Fellow, IEEE*, Rongke Liu, *Member, IEEE*, Majid Safari, *Member, IEEE*, and Pan Cao, *Member, IEEE*

***Abstract*— Channel asymmetry and channel correlation are frequently encountered in wireless communication systems. Orthogonal transmission schemes are usually inefficient in dealing with these problems. In this paper, in order to boost the throughput performance for multiple-input multiple-output broadcast communications in the presence of channel asymmetry and/or channel correlation, we study optimization algorithms for multi-user superposition coding beamforming (SCBF). Starting with solving the minimum power optimization problem for the two-user case, we derive the optimal solution structure of the problem and two types of dedicated algorithms that could efficiently find the optimal solutions with all parameter setups. Extensions are then made to the same problem with the signals of more than two users multiplexed in the power domain as well as to the rate region computation problem. Finally, to adapt our algorithms to more general cases, novel hybrid precoding schemes are proposed, where certain user grouping strategy is used to combine zero-forcing beamforming and SCBF. Numerical simulations are provided to show that with our algorithms, a considerable performance gain is achieved by SCBF compared to the other orthogonal transmission methods.**

***Index Terms*— Optimizations, beamforming, superposition coding, channel correlations, second-order conic programs.**

I. Introduction

TRANSMIT beamforming or linear precoding at the transmitter side plays an important role in striking a balance between high efficiency and low complexity in a multi-user MIMO (MU-MIMO) wireless communication system. The capacity approaching method dirty paper coding (DPC) [1], and some other non-linear precoding techniques [2], [3] usually can achieve better performance than transmit beamforming, but require high computational complexity that makes them less attractive in practice.

One disadvantage of linear beamforming in comparison with nonlinear methods is that its efficiency is undermined in ill-conditioned channels. As shown in [4], through regularizations, linear precoding could achieve performance that is close to the DPC capacity bound in the low signal-to-noise ratio (SNR) region. However, the performance gap becomes prominent at higher SNRs. To boost system performance while preserving the low complexity characteristic of linear precoding, [5] and [6] propose user grouping and user selection methods respectively for the downlink of cellular networks. Using zero-forcing beamforming (ZFBF), [5] exploits user diversity and devises a user selection policy by evaluating the orthogonality between the channels. Reference [6] uses DPC to deal with the interference between the group members and the low-complexity ZFBF to eliminate the inter group interference.

Two factors are important in creating an ill-conditioned MU-MIMO channel. One concerns the correlation between the sub-channels while the other is related with channel asymmetry, which may be caused by either heterogeneous receivers or the large variability of the channel conditions. The combination of beamforming and superposition coding, denoted as SCBF hereafter, provides an efficient way to deal with these two factors. As shown later in our work, SCBF provides robustness to channel correlations and achieves superior throughput performance in an asymmetric broadcast channel. SCBF multiplexes users' signals in the power domain and implements successive interference cancellation (SIC) with a pre-defined order at the user side. The implementation complexity of SCBF is low at the transmitter since it only employs linear precoding. At the user end, additional computation is needed due to SIC but remains acceptable using state of the art digital signal processing (DSP) techniques [9].

Existing research studying the combination of superposition coding and MIMO communications can be found in [10]–[17], and [19]. Due to the high complexity of the studied problems, heuristic suboptimal algorithms are usually utilized but the literature is not yet clear about how to combine superposition coding and beamforming in an efficient way. In order to exploit the full potential of SCBF, we focus on the design of efficient optimization algorithms for SCBF in this paper. We start by studying the minimum power optimization problem. Notice that [15] also studies a similar scenario and applies SCBF to

Manuscript received July 12, 2017; revised January 16, 2018 and June 12, 2018; accepted August 12, 2018. This work was supported by the National Natural Science Foundation of China under Grant 91438116. This paper was presented in part at the IEEE ICC 2017 1st International Workshop on Satellite Communications. The associate editor coordinating the review of this paper and approving it for publication was C. R. Murthy. *(Corresponding authors: John S. Thompson; Rongke Liu.)*

X. Shi is with the Institute for Digital Communications, University of Edinburgh, Edinburgh EH3 9JL, U.K., and also with Beihang University, Beijing 100191, China (e-mail: xiaoyan.shi@ed.ac.uk).

J. S. Thompson and M. Safari are with the Institute for Digital Communications, University of Edinburgh, Edinburgh EH3 9JL, U.K. (e-mail: john.thompson@ed.ac.uk; majid.safari@ed.ac.uk).

R. Liu is with the School of Electronics and Information Engineering, Beihang University, Beijing 100191, China (e-mail: rongke_liu@buaa.edu.cn).

P. Cao is with the School of Engineering and Technology, University of Hertfordshire, Hertfordshire AL10 9AB, U.K. (e-mail: p.cao@herts.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2018.2866580

multi-resolution broadcast communications. However, due to the non-convexity of the problem, the optimality of the algorithm in [15] is not guaranteed. Reference [16] does find the optimal solution for the same problem, but the solution can be computed with low complexity only when certain conditions are met.

Compared to the existing work, this paper has three main contributions. First, instead of solving the dual problem, we explicitly transform the original problem into a second-order conic program (SOCP) and derive the structure of the optimal solution for the original problem. Reference [17] also solves the primal problem directly by using the optimality conditions but only for special cases where one of the constraints related to the primal problem is assumed inactive. For larger networks with more than two users, [17] proposes a user pairing algorithm and serves each pair of users in a different time or frequency block. We propose a novel hybrid precoding scheme that is based on user grouping and therefore all users are served simultaneously. Numerical simulations show that our method achieves much better performance than [17]. Second, by exploiting the structure of the optimal solution, we derive a fixed point iterative algorithm and another water-filling type algorithm, which has lower computational complexity, to find the optimal solution for the two-user minimum power beamforming problem. Using similar method, we also find the optimal solution to the rate maximization problem and therefore compute the 2-user region achieved by SCBF. Last but not the least, our algorithms can be conveniently generalized to cases where more than two layers of information are multiplexed in the power domain. We emphasize that in most other research, although methods are designed for large number of users, only two layers of information are assumed to be multiplexed in power domain and be separated at the stronger users using SIC. In contrast, our work provides an efficient algorithm that supports arbitrary layers of information. In particular for the 3-layer case, we devise algorithms that in most practical scenarios yield solutions that stay very close to the optimal ones, with computational complexity similar to the best linear beamforming [7].

The paper is outlined as follows. In Section II, the system model and the problem formulation are given. In Section III, the structure of the optimal solution is analyzed and three types of optimality achieving algorithms for the two user problem are derived. In Section IV, we explain how to generalize our method to cases where more than two users are involved as well as to the rate region computation problem. Finally in Section V, numerical simulations are given and we perform a detailed analysis of SCBF with Rician fading channels.

Notation: $\|\cdot\|$ represents vector ℓ_2-norm, whereas $\mathbf{I}$ denotes the diagonal matrix with all one diagonal elements. $\mathrm{Diag}(\cdot)$ translates the enclosed vector or elements into a diagonal matrix. For any matrix $\mathbf{A}$, $\mathrm{Col}(\mathbf{A})$ denotes the column space of the matrix and $\mathcal{A}^{\perp}$ denotes the orthogonal vector space of $\mathrm{Col}(\mathbf{A})$. For a complex scalar r, r^* denotes the complex conjugate. $\mathbf{0}$ is the zero vector.

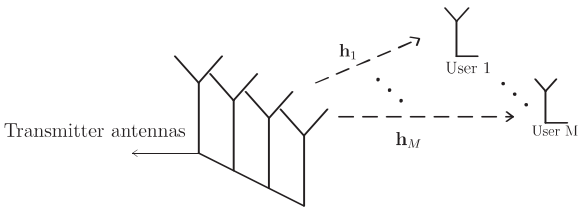


Fig. 1. An illustration of the system model for SCBF in downlink MU-MIMO communications.

II. System Model and Problem Formulation

A. Downlink Multi-User MIMO Precoding With Superposition Coding

The communication scenario of this paper can be modeled as a MU-MIMO channel where the transmitter is equipped with N antennas and each user is equipped with one antenna. We denote the channel by an N by M complex matrix $\mathbf{H} \in \mathbb{C}^{N\times M}$ (a number of M users is assumed), the i^{th} column of which, $\mathbf{h}_i$, represents the sub-channel vector that User i sees. We assume a frequency-flat and block fading channel, i.e. the channel gain $\mathbf{h}_i$ stays constant within the duration of a codeword transmission. The transmitter is assumed to have accurate channel side information. To model transmit beamforming, the transmitted signal is denoted as

$$\mathbf{s} = \sum_{i=1}^{M} s_i \mathbf{w}_i, \tag{1}$$

where s_i is a complex scalar representing the information symbol of User i, $\mathbf{w}_i$ is the associated beamforming vector of size N. The received signal at User i could be represented as

$$y_i = \mathbf{h}_i^H \mathbf{s} + z_i = \mathbf{h}_i^H \sum_{k=1}^{M} s_k \mathbf{w}_k + z_i, \quad 1 \le i \le M, \tag{2}$$

where z_i is the zero-mean complex additive white Gaussian noise (AWGN) at User i with variance σ_i^2. We assume that each user is interested in its own data stream. However, it should be clear that the algorithms we study here should be equally effective for the broadcast case.

To implement SIC at the user end, a decoding sequence should be established, which is related to the property of the channel degradedness in broadcast channel [9]. First, we consider Rayleigh fading channels, i.e. $\mathbf{h}_i \sim \mathcal{CN}(0, (1/(d_i^{\eta}N))\mathbf{I})$, where d_i is the distance between the i^{th} user and the base station (BS), η the path loss exponent, set to be 2 throughout this paper, and N the number of transmit antennas. The users can be ordered by considering only their large-scale path-loss attenuation, which is related to their distance to the BS. Another type of channel degradedness exists in Rician fading MIMO channels. Rician fading models both the LOS and the scattered signals, i.e $\mathbf{H} = (\mathbf{h}_1, \mathbf{h}_2) = \mathbf{H}_{los}\mathbf{G}_l + \mathbf{H}_{nlos}\mathbf{G}_n$ if a two user case is considered, where

$$\mathbf{G}_l = \mathrm{Diag}\left(\sqrt{\frac{v_k}{v_k+1}}\right), \quad \mathbf{G}_n = \mathrm{Diag}\left(\sqrt{\frac{1}{v_k+1}}\right), \; k = 1, 2. \tag{3}$$

We use v_k to denote the Rician K-factors. The entries of $\mathbf{H}_{nlos}$ are independent and identically distributed (i.i.d.) as $\mathcal{CN}(0,1)$. A uniform linear array (ULA) [21] is assumed, i.e. $\mathbf{h}_{i,los} = [1, e^{-j2\pi\frac{d}{\lambda}\cos\theta_i}, \ldots, e^{-j2\pi(N-1)\frac{d}{\lambda}\cos\theta_i}]^T$, where d and λ is the distance between the neighboring elements and the wavelength respectively. The scalar $\theta_i \in [0, +\pi]$ is the angle of arrival (AoA). We will discuss the application of superposition coding to the Rician MIMO channel in more detail in Section V.

For any given channel condition, an ordered set $\mathcal{S} = \{u_1, u_2, \ldots, u_M\}$ can then be used to represent both the entire users as well as the decoding sequence. A user with a larger index is considered to have stronger channel. Note that u_i can always decode the information sent to u_j for any $j \leq i$ under the following conditions

$$\min_i(\text{SINR}_i^j) \geq \gamma_j, \quad j \leq i \leq M, \ 1 \leq i \leq M, \tag{4}$$

where γ_i is a positive constant and SINR_i^j is the signal-to-interference-plus-noise ratio (SINR) for u_i to decode u_j's information. Due to the effect of SIC, we can also write that

$$\text{SINR}_i^j = \frac{|\mathbf{h}_i^H \mathbf{w}_j|^2}{\sum_{m=j+1}^{M} |\mathbf{h}_i^H \mathbf{w}_m|^2 + \sigma_i^2}, \quad 1 \leq i \leq M, \ j \leq i, \tag{5}$$

and the achievable rate for u_i, denoted as R_i is related to γ_i by $2^{R_i} - 1 = \gamma_i$.

B. Optimization Problem Formulation

In this section, formal representations are provided for the two types of optimization problems that we are interested in this paper.

1) Minimum Power Beamforming Optimization: We will first study the problem where the beamforming vectors are optimized to minimize the total transmission power with a fixed QoS constraint upon each user, as represented in (6). The problem is denoted as the minimum power beamforming optimization and has received much attention in recent research such as [15]–[18].

$$\min_{\mathbf{w}_i} \ \sum_{i=1}^{M} \mathbf{w}_i^H \mathbf{w}_i \tag{6a}$$

$$\text{s.t.} \ \min_k(\text{SINR}_k^i) \geq \gamma_i, \quad i \leq k \leq M, \ 1 \leq i \leq M, \tag{6b}$$

It can be observed that although the objective function of problem (6) is convex, the constraints are not because of the quadratic terms on the left side of the inequalities.

2) Rate Region Computation: To provide more insight into the characteristics of the combination of beamforming and superposition coding, we are also interested in the two user rate region computation problem, denoted as follows

$$\max_{\mathbf{w}_i, \gamma_2} \ \gamma_2 \tag{7a}$$

$$\text{s.t.} \ \sum_{i=1}^{M} \mathbf{w}_i^H \mathbf{w}_i \leq P, \tag{7b}$$

$$\min_{k \in 1,2}(\text{SINR}_k^1) \geq \gamma_1, \tag{7c}$$

$$\text{SINR}_2^2 \geq \gamma_2, \tag{7d}$$

where P is the upper limit for the total transmission power. The optimization objective is to find the optimal rate for the stronger user, i.e. User 2, with a lower rate bound (equivalently represented by the constant SINR threshold γ_1) on User 1.

III. Minimum Power Non-Orthogonal Beamforming Optimization

We start with studying the two-user minimum power optimization problem, i.e. (6) with $M = 2$. According to our prior analysis on the decoding sequence, u_1 is considered as the 'weaker user' and its information should be decoded at both receivers. To ease our analysis, we explicitly write this special case as follows

$$\min_{\mathbf{w_1}, \mathbf{w_2}} \ \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \tag{8a}$$

$$\text{s.t.} \ \gamma_1 \left(|\mathbf{h}_1^H \mathbf{w}_2|^2 + \sigma_1^2\right) \leq |\mathbf{h}_1^H \mathbf{w}_1|^2, \tag{8b}$$

$$\gamma_1 \left(|\mathbf{h}_2^H \mathbf{w}_2|^2 + \sigma_2^2\right) \leq |\mathbf{h}_2^H \mathbf{w}_1|^2, \tag{8c}$$

$$\gamma_2 \sigma_2^2 \leq |\mathbf{h}_2^H \mathbf{w}_2|^2. \tag{8d}$$

We propose a novel method that explicitly transforms this original problem into an SOCP. A fixed-point iterative algorithm as well as a novel water-filling type algorithm are proposed to produce optimal solutions with low computational complexity. Constraints similar to (8b) and (8c) are also frequently met in other beamforming optimizations, such as multi-group multicast beamforming where a single channel signature $\mathbf{w}_1$ serves multiple downlink users [25].

A. Conic Optimization Solution

First the beamforming vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in (8) are rotationally equivalent, i.e. for any $\theta_i \in [0, 2\pi]$, after replacing the solution of $\mathbf{w}_i$ with $e^{j\theta_i}\mathbf{w}_i$, the values of the terms in inequalities (8b)-(8d) and the objective function remain unchanged. Therefore (8d) can be transformed into $\sigma_2 \leq \frac{1}{\sqrt{\gamma_2}}\Re(\mathbf{h}_2^H \mathbf{w}_2)$, which is a second-order conic constraint.

However, it is not straightforward to apply such equivalent transformation to constraints (8b) and (8c) at the same time. The difficulty lies in the fact that unless $\mathbf{h}_1$ and $\mathbf{h}_2$ are linearly dependent, merely rotating $\mathbf{w}_1$ is not sufficient to make both $\mathbf{h}_1^H \mathbf{w}_1$ and $\mathbf{h}_2^H \mathbf{w}_1$ non-negative at the same time. A rotational coefficient is further required to be explicitly multiplied with one of the term. To determine this coefficient, suppose that the QR decomposition of $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2]$ is

$$\mathbf{H} = \mathbf{V}\mathbf{R} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}, \tag{9}$$

where $\mathbf{R}$ is a 2 by 2 upper triangular matrix with non-negative diagonal entries, i.e. $r_{11} \geq 0, r_{22} \geq 0$. The columns of $\mathbf{V}$ form a part of an orthonormal basis of N-dimensional complex space. We impose the following fixed structure upon $\mathbf{w}_1$

$$\mathbf{w}_1 = K_1 \frac{r_{12}}{|r_{12}|^2}\mathbf{v}_1 + K_2 \frac{r_{22}}{|r_{22}|^2}\mathbf{v}_2, \tag{10}$$

where K_1, K_2 are positive real scalars. As later proved in Proposition 1, the optimality of the original problem can be achieved even when the fixed structure of (10) is assumed. Check that with (10), $\mathbf{h}_2^H \mathbf{w}_1$ will always be a positive scalar.

In order to make $\mathbf{h}_1^H\mathbf{w}_1$ positive, we multiply $\mathbf{h}_1$ with a rotation coefficient and make the result replace $\mathbf{h}_1$ as the new channel gain, i.e. $\hat{\mathbf{h}_1} = \mathbf{h}_1 e^{j\phi_{12}}$, where ϕ_{12} is the phase of $\mathbf{h}_1^H\mathbf{h}_2$. Now we can write that

$$\hat{\mathbf{h}_1}^H\mathbf{w}_1 = r_{11}K_1\frac{r_{12}}{|r_{12}|^2}e^{-j\phi_{12}}. \tag{11}$$

Since $r_{12}e^{-j\phi_{12}} = \mathbf{h}_1^H\mathbf{h}_2\, e^{-j\phi_{12}}/r_{11}$ and r_{11} are both positive reals, we conclude that $\hat{\mathbf{h}_1}^H\mathbf{w}_1$ is also a positive real. Rewrite problem (8) as

$$\min_{\mathbf{w_1},\mathbf{w_2}} \ \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 * \tag{12a}$$

$$\text{s.t. } \frac{1}{\sqrt{\gamma_1}}\Re(\hat{\mathbf{h}_1}^H\mathbf{w}_1) \geq \sqrt{|\mathbf{h}_1^H\mathbf{w}_2|^2 + \sigma_1^2}, \tag{12b}$$

$$\frac{1}{\sqrt{\gamma_1}}\Re(\mathbf{h}_2^H\mathbf{w}_1) \geq \sqrt{|\mathbf{h}_2^H\mathbf{w}_2|^2 + \sigma_2^2}, \tag{12c}$$

$$\frac{1}{\sqrt{\gamma_2}}\Re(\mathbf{h}_2^H\mathbf{w}_2) \geq \sigma_2. \tag{12d}$$

Since all of the constraints are of second-order conic type, we could solve (12) by using SOCP optimization tools and more importantly an optimal solution can be found. Since the feasible set of (12) could be different from (8), we now prove that the same optimal objective value can actually be achieved by solving either one of the problems.

Proposition 1: Suppose that the optimal solutions of problem (8) *and problem* (12) *are* $[\mathbf{w}_{\text{o1}}, \mathbf{w}_{\text{o2}}]$ *and* $[\mathbf{w}_{\text{p1}}, \mathbf{w}_{\text{p2}}]$ *respectively. Then* $\|\mathbf{w}_{\text{o1}}\|^2 + \|\mathbf{w}_{\text{o2}}\|^2 = \|\mathbf{w}_{\text{p1}}\|^2 + \|\mathbf{w}_{\text{p2}}\|^2$ *holds, i.e. the two problems achieve the same objective value via their own optimal solutions.*

Proof: First, without loss of optimality, we could represent the optimal solution of $\mathbf{w}_1$ as $\mathbf{w}_{\text{o1}} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$, where α_1, α_2 are complex scalars. Then using (9), we could derive

$$|\mathbf{h}_1^{\text{H}}\mathbf{w}_{\text{o1}}| = |(r_{11}\mathbf{v}_1)^H(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2)| = |r_{11}^*\alpha_1|, \tag{13}$$

$$|\mathbf{h}_2^{\text{H}}\mathbf{w}_{\text{o1}}| = |(r_{12}\mathbf{v}_1 + r_{22}\mathbf{v}_2)^H(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2)|. \tag{14}$$

Accordingly, by using (10), which is the structure of the solutions to (12), we could derive

$$|\hat{\mathbf{h}_1}^H\mathbf{w}_{\text{p1}}| = |\mathbf{h}_1^H\mathbf{w}_{\text{p1}}| = |r_{11}^*K_1\frac{r_{12}}{|r_{12}|^2}|, \tag{15}$$

$$|\mathbf{h}_2^H\mathbf{w}_{\text{p1}}| = K_1 + K_2. \tag{16}$$

To guarantee $|\mathbf{h}_1^{\text{H}}\mathbf{w}_{\text{o1}}| = |\hat{\mathbf{h}_1}^{\text{H}}\mathbf{w}_{\text{p1}}|$, we would have $K_1 = |r_{12}^*\alpha_1|$. When $K_2 = |r_{22}^*\alpha_2|$, (17) and (18) are obtained as

$$K_1 + K_2 = |r_{12}^*\alpha_1| + |r_{22}^*\alpha_2| \overset{(a)}{\geq} |r_{12}^*\alpha_1 + r_{22}^*\alpha_2|. \tag{17}$$

$$\|\mathbf{w}_{\text{p1}}\|^2 = \left|K_1\frac{r_{12}}{|r_{12}|^2}\right|^2 + \left|K_2\frac{r_{22}}{|r_{22}|^2}\right|^2 = \|\mathbf{w}_{\text{o1}}\|^2 \tag{18}$$

We now use proof by contradiction to show that the relationship (a) in (17) can only be satisfied with equality. Suppose that the inequality is strict. Then to have $K_1 + K_2 = |r_{12}^*\alpha_1 + r_{22}^*\alpha_2|$, K_2 must become smaller than $|r_{22}^*\alpha_2|$ if $K_1 = |r_{12}^*\alpha_1|$. Due to (18), this would further indicate that, $\|\mathbf{w}_{\text{p1}}\|^2 < \|\mathbf{w}_{\text{o1}}\|^2$. Furthermore, to satisfy $|\mathbf{h}_1^{\text{H}}\mathbf{w}_{\text{o2}}| = |\hat{\mathbf{h}_1}^{\text{H}}\mathbf{w}_{\text{p2}}|$ and $|\mathbf{h}_2^{\text{H}}\mathbf{w}_{\text{o2}}| = |\mathbf{h}_2^{\text{H}}\mathbf{w}_{\text{p2}}|$ at the same time, the difference between $\mathbf{w}_{\text{o2}}$ and $\mathbf{w}_{\text{p2}}$ could simply be a rotation coefficient. Therefore, we would have $\|\mathbf{w}_{\text{p1}}\|^2 + \|\mathbf{w}_{\text{p2}}\|^2 < \|\mathbf{w}_{\text{o1}}\|^2 + \|\mathbf{w}_{\text{o2}}\|^2$, which is contradicted by the fact that $\mathbf{w}_{\text{o1}}, \mathbf{w}_{\text{o2}}$ are the optimal solution of the original problem.

In conclusion, we can construct one solution of (12) that achieves the same objective value as the one achieved by the optimal solution of (8). Since the feasible set of (12) is also included by the feasible set of its original problem (8), we finally prove Proposition 1. □

B. Solution Through Optimality Conditions

In this section, we deal with the problem of (8) by checking the Karush-Kuhn-Tucker (KKT) conditions. We highlight that [16] or [17] provides a low-complexity algorithm only for some special cases, while in this paper, by exploiting the structure of optimal solutions, we derive optimality achieving algorithms with low computational complexity for all scenarios.

As shown in [20], for problem (8), the KKT conditions are both necessary and sufficient to achieve optimality under the condition that (8) is strictly feasible. For this required strict feasibility, it could be seen that as long as the channel gains are linearly independent, there is always a solution of (8) that satisfies the constraints strictly.

First, the Lagrangian function of problem (8) can be written as

$$\mathfrak{L} = \sum_{i=1}^{3}\lambda_i\sigma_i^2 + \mathbf{w}_1^H\left(\mathbf{I} - \frac{\lambda_1}{\gamma_1}\mathbf{h}_1\mathbf{h}_1^H - \frac{\lambda_2}{\gamma_1}\mathbf{h}_2\mathbf{h}_2^H\right)\mathbf{w}_1 + \mathbf{w}_2^H\left(\mathbf{I} + \lambda_1\mathbf{h}_1\mathbf{h}_1^H + \lambda_2\mathbf{h}_2\mathbf{h}_2^H - \frac{\lambda_3}{\gamma_2}\right)\mathbf{w}_2, \tag{19}$$

where $\lambda_i \geq 0, i = 1,2,3$ are the three dual variables that correspond to constraints (8b), (8c) and (8d) respectively. According to previous analysis, when $\mathbf{h}_1$ and $\mathbf{h}_2$ are linearly independent, the following conditions should be both sufficient and necessary to achieve optimality:

1) *Feasibility*: Inequalities (8b),(8c),(8d)
2) *Complementary Slackness*: For $i = 1,2,3$, either $\lambda_i = 0$ or the corresponding constraints should be satisfied with equality.
3) *Zero Derivatives*: Since the derivatives of the Lagrange function with respect to (w.r.t.) the dual variables equal to zero, we have

$$\frac{\partial\mathfrak{L}}{\partial\mathbf{w}_1} = \left(\mathbf{I} - \frac{\lambda_1}{\gamma_1}\mathbf{h}_1\mathbf{h}_1^H - \frac{\lambda_2}{\gamma_1}\mathbf{h}_2\mathbf{h}_2^H\right)\mathbf{w}_1 = 0, \tag{20}$$

$$\frac{\partial\mathfrak{L}}{\partial\mathbf{w}_2} = \left(\mathbf{I} + \lambda_1\mathbf{h}_1\mathbf{h}_1^H + \lambda_2\mathbf{h}_2\mathbf{h}_2^H - \frac{\lambda_3}{\gamma_2}\mathbf{h}_2\mathbf{h}_2^H\right)\mathbf{w}_2 = 0. \tag{21}$$

We can rewrite conditions (20) and (21) as follows

$$\mathbf{w}_1 = \frac{\lambda_1}{\gamma_1}\mathbf{h}_1\mathbf{h}_1^H\mathbf{w}_1 + \frac{\lambda_2}{\gamma_1}\mathbf{h}_2\mathbf{h}_2^H\mathbf{w}_1, \tag{22}$$

$$\mathbf{w}_2 = \left(\mathbf{I} + \lambda_1\mathbf{h}_1\mathbf{h}_1^H + \lambda_2\mathbf{h}_2\mathbf{h}_2^H\right)^{-1}\frac{\lambda_3}{\gamma_2}\mathbf{h}_2\mathbf{h}_2^H\mathbf{w}_2. \tag{23}$$

We define $\mathbf{G}$ as $\left(\mathbf{I} + \lambda_1\mathbf{h}_1\mathbf{h}_1^H + \lambda_2\mathbf{h}_2\mathbf{h}_2^H\right)$. Next we show that we could calculate the dual variables in (20) and (21) through

solving a matrix equation. Inspired by Proposition 1 where we impose a fixed structure upon the solution without loss of optimality, we also assume (10) in the following derivation. We normalize the channel gain w.r.t. the noise for convenience so that $\mathbf{z}_i$ has unit variance, i.e. $\sigma_1^2 = \sigma_2^2 = 1$.

First, notice that at the optimal point the constraint (8d) should actually be satisfied with equality. This follows from the same argument as the ones made in typical orthogonal beamforming optimizations [18]. However, the equality may not hold for the other two constraints in problem (8) at the optimal point. Therefore due to complementary slackness, only λ_3 is guaranteed to be strictly positive and a case by case analysis should be made after taking the positivity of λ_1 and λ_2 into consideration. We start with the case where all dual variables are strictly positive. First by combining (10) with (22) and computing $\left|\mathbf{h}_2^H\mathbf{w}_1\right|^2$, we could derive

$$\left|\mathbf{h}_2^H\mathbf{w}_1\right|^2 = \left|\mathbf{h}_2^H\left(\frac{\lambda_1}{\gamma_1}\mathbf{h}_1\mathbf{h}_1^H\mathbf{w}_1 + \frac{\lambda_2}{\gamma_1}\mathbf{h}_2\mathbf{h}_2^H\mathbf{w}_1\right)\right|^2 \tag{24}$$

$$= \left|\frac{\lambda_1}{\gamma_1}\left|\mathbf{h}_2^H\mathbf{h}_1\right|^2\frac{K_1}{|r_{12}|^2} + \frac{\lambda_2}{\gamma_1}\|\mathbf{h}_2\|^2(K_1+K_2)\right|^2$$

$$= |K_1+K_2|^2, \tag{25}$$

where the second equation in (25) is obtained by combining (9) and (10). After taking the root square of (25), we obtain the following equation

$$\frac{\lambda_1}{\gamma_1}|\mathbf{h}_2^H\mathbf{h}_1|^2\frac{1}{|r_{12}|^2}K_1 + \frac{\lambda_2}{\gamma_1}\|\mathbf{h}_2\|^2(K_1+K_2) = K_1+K_2. \tag{26}$$

Similarly, through computing $\left|\mathbf{h}_1^H\mathbf{w}_1\right|^2$, we could also derive

$$\frac{\lambda_1}{\gamma_1}\|\mathbf{h}_1\|^2\frac{K_1}{|r_{12}|^2} + \frac{\lambda_2}{\gamma_1}(K_1+K_2) = \frac{1}{|r_{12}|^2}K_1. \tag{27}$$

Next by using (21), we obtain another equation indicating how to ascertain λ_3 when the other two dual variables are found

$$\mathbf{h}_2^H\mathbf{G}^{-1}\mathbf{h}_2\frac{\lambda_3}{\gamma_2} = 1. \tag{28}$$

Through (8c) and (8d), we further have $\mathbf{h}_2^H\mathbf{w}_1 = K_1+K_2 = \sqrt{\gamma_1(\gamma_2+1)}$. Finally through (8b) and (27), we derive

$$|r_{11}^*r_{12}|\left(\frac{\lambda_1}{\gamma_1}\|\mathbf{h}_1\|^2\frac{K_1}{|r_{12}|^2} + \frac{\lambda_2}{\gamma_1}(K_1+K_2)\right)$$

$$= \sqrt{\gamma_1\gamma_2\frac{|\mathbf{h}_1^H\mathbf{G}^{-1}\mathbf{h}_2|}{|\mathbf{h}_2^H\mathbf{G}^{-1}\mathbf{h}_2|} + \gamma_1}. \tag{29}$$

Now that we have five independent equations (26)-(29), it becomes possible to find the five unknown variables, including the three dual variables and the coefficients in (10), i.e. K_1 and K_2. In a similar way, we deal with the case when $\lambda_1 = 0$ or $\lambda_2 = 0$. All of these cases are covered in Algorithm 1 with $c_1 = \frac{a_{11}\|\mathbf{h}_1\|}{\gamma_1|r_{12}|}$, $c_2 = \frac{\lambda_1|a_{12}|^2}{\gamma_1|r_{12}|^2}$, $c_3 = \frac{R_c}{\gamma_1}(a_{12} - a_{12}^{-1}a_{11}a_{22})$, where a_{ij} is the absolute value of the inner product between $\mathbf{h}_i$ and $\mathbf{h}_j$, $R_c = (\gamma_1(\gamma_2+1))^{0.5}$.

A detailed analysis of the convergence properties for Stage-B in Algorithm 1 is provided in VI and we conclude the result in the following proposition.

Algorithm 1 Minimum Power SCBF With Two Users

Stage-A: For $i = 1, 2$,

1: Compute: $\lambda_i = \frac{\gamma_1}{\|\mathbf{h}_i\|^2}$, $\mathbf{G}^{-1} = (\mathbf{I} + \lambda_i\mathbf{h}_i\mathbf{h}_i^H)^{-1}$.

2: If $(i == 1)$ $K_1 = \left(\gamma_1\gamma_2\frac{|\mathbf{h}_1^H\mathbf{G}^{-1}\mathbf{h}_2|^2}{|\mathbf{h}_2^H\mathbf{G}^{-1}\mathbf{h}_2|^2} + \gamma_1\right)^{0.5} / (c_1\lambda_1)$, $K_2 = (c_2\lambda_1 - 1)K_1$.

3: Else $K_1 = \frac{1}{\gamma_1}\sqrt{\gamma_1(\gamma_2+1)}\|r_{12}\|^2\lambda_2$, $K_2 = \sqrt{\gamma_1(\gamma_2+1)} - K_1$.

4: Check if the KKT conditions are not satisfied. If not, go to Stage-B.

Stage-B: Initialize by randomly generating λ_1^0, λ_2^0 and set t:=0. Repeat the following procedures until convergence or exceeding the required number of iterations

1: Compute the following terms

$$\mathbf{G}^{-1} = (\mathbf{I} + \lambda_1^t\mathbf{h}_1\mathbf{h}_1^H + \lambda_2^t\mathbf{h}_2\mathbf{h}_2^H)^{-1}, \tag{30}$$

$$\lambda_2^{t+1} = \frac{1}{c_3}\left[(\gamma_1\gamma_2\frac{|\mathbf{h}_1^H\mathbf{G}^{-1}\mathbf{h}_2|^2}{|\mathbf{h}_2^H\mathbf{G}^{-1}\mathbf{h}_2|^2} + \gamma_1)^{0.5} - R_c\frac{a_{11}}{a_{12}}\right]. \tag{31}$$

2: Solve the following linear equation set.

$$\frac{a_{12}R_c}{\gamma_1}\lambda_2^{t+1} = -\frac{a_{11}}{\gamma_1}\frac{\|\mathbf{h}_1\|}{|r_{12}|}x + \frac{\|\mathbf{h}_1\|}{|r_{12}|}y, \tag{32}$$

$$\frac{a_{22}R_c}{\gamma_1}\lambda_2^{t+1} = -\frac{a_{12}}{\gamma_1}\frac{\|\mathbf{h}_1\|}{|r_{12}|}x + R_c. \tag{33}$$

3: Update the dual variables with the obtained solutions x_o and y_o: $\lambda_1^{t+1} = x_o/y_o$, $t = t+1$.

Proposition 2: The fixed point for the set of functions (30)-(33) *is denoted as* $(\lambda_1^f, \lambda_2^f)$. *Given the initial point* $(\lambda_1^0, \lambda_2^0)$, *there exists a convergence area* $\mathcal{N}$, *such that if* $(\lambda_1^0, \lambda_2^0), (\lambda_1^f, \lambda_2^f) \in \mathcal{N}$, *the iteration process defined by functions* (30)-(33) *will converge to the fixed point.*

Proof: Please see A. □

Due to the requirement for constraining the first derivative of the fixed point function, derived by merging (A.4) and (A.5), to be less than one, the initial point for the iteration should lie within a certain range to guarantee convergence (for details please check Appendix A). Therefore, only a local convergence result is obtained. To tackle this problem, we propose a novel incremental optimization method which is globally convergent, providing an alternative to Algorithm 1.

C. Incremental Optimization Algorithm

First we define problem (8) with constraint (8c) removed as the degraded problem, denoted as $\mathbf{P}_D$, to the original problem (8). For $\mathbf{P}_D$, the optimal solution of $\mathbf{w}_1$, denoted as $\mathbf{w}_{1\text{od}}$, should only contain one component along the $\mathbf{h}_1$. According to (10), we can write explicitly that

$$\mathbf{w}_{1\text{od}} = K_{1\text{od}}\frac{r_{12}}{|r_{12}|^2}\mathbf{v}_1 + K_{2\text{od}}\frac{r_{22}}{|r_{22}|^2}\mathbf{v}_2, \tag{34}$$

where $K_{1\text{od}} > 0$ and $K_{2\text{od}} = 0$. Similarly, we use $\mathbf{w}_{2\text{od}}$ to denote the optimal solution of $\mathbf{w}_2$ to problem $\mathbf{P}_D$ and without

TABLE I
EVENT ϵ_{ND} IN RAYLEIGH FADING: $d_1 = 1$m, $N = 5$, $\mathrm{R}_1 = \mathrm{R}_2 = 2$ bit/s/Hz

d_2/m	5	10	15
P(ϵ_{ND})	92%	65%	33%
$E(\zeta)$	-0.04	0.04	0.03
RMSD(ζ)	0.35	0.12	0.05

loss of optimality we can write

$$\mathbf{w}_2 = K_{21}\frac{r_{12}}{|r_{12}|^2}\mathbf{v}_1 + K_{22}\frac{r_{22}}{|r_{22}|^2}\mathbf{v}_2, \tag{35}$$

and $K_{21\mathrm{od}}$ and $K_{22\mathrm{od}}$ are used to denote the corresponding optimal values for the degraded problem $\mathbf{P}_D$. Now add constraint (8c) back to $\mathbf{P}_D$ and the original problem is restored. If constraint (8c) is satisfied by the solution to $\mathbf{P}_D$, then the original problem has already been solved at the same time when we implement Stage-A in Algorithm 1 to solve problem $\mathbf{P}_D$. The next two parts of this section will first show that in practice, there can be high probability that the solution to $\mathbf{P}_D$ fails to satisfy (8c). We define such a scenario as the non-degraded case. For this case, we discuss below how to find the optimal solution for the original problem. Meanwhile, we also propose to combine SCBF with ZFBF to further enhance the robustness of our algorithm to cases where channel asymmetry is less prominent.

1) Non-Degraded Case and the Combination of SCBF With ZFBF: We note that in [16], an efficient algorithm has also been given for the degraded problem. However, when constraint (8c) is not satisfied by the solution of $\mathbf{P}_D$, [16] fails to provide efficient algorithm to achieve optimality of the original problem. Let us use ϵ_{ND} to denote the event where the optimal solution of $\mathbf{P}_D$ is not equal to that of the original problem, i.e the non-degraded case. To characterize the frequency of the occurrence of ϵ_{ND} in Rayleigh fading channels, we take $\mathrm{N_{tot}} = 1000$ independent channel realizations and compute probability of occurrence of ϵ_{ND}. We define $\mathrm{P_{zi}}$ as the total transmission power needed for the i^{th} realization of channel when ZFBF is adopted and $\mathrm{P_{si}}$ is the corresponding value when SCBF is used. We define $\zeta_i = (\mathrm{P_{zi}} - \mathrm{P_{si}})/\mathrm{P_{zi}}$. Here using the notations in (8), the ZFBF beamforming vectors satisfy

$$\begin{aligned}(\mathbf{h}_1^H\mathbf{w}_1)^2 &= \gamma_1\sigma_1^2 = \gamma_1^{\mathrm{ZFBF}}, \quad (\mathbf{h}_2^H\mathbf{w}_1) = 0,\\ (\mathbf{h}_2^H\mathbf{w}_2)^2 &= \gamma_2\sigma_2^2 = \gamma_2^{\mathrm{ZFBF}}, \quad (\mathbf{h}_1^H\mathbf{w}_2) = 0,\end{aligned} \tag{36}$$

and can be computed according to

$$\mathbf{w}_1 = \frac{\sqrt{\gamma_1^{\mathrm{ZFBF}}}\mathbf{a}_1}{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2\sin^2\theta}, \quad \mathbf{w}_2 = \frac{\sqrt{\gamma_2^{\mathrm{ZFBF}}}\mathbf{a}_2}{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2\sin^2\theta}, \tag{37}$$

where $\mathbf{a}_1 = \|\mathbf{h}_2\|^2\mathbf{h}_1 - (\mathbf{h}_2^H\mathbf{h}_1)\mathbf{h}_2, \mathbf{a}_2 = \|\mathbf{h}_1\|^2\mathbf{h}_2 - (\mathbf{h}_1^H\mathbf{h}_2)\mathbf{h}_1$ and $\cos^2\theta = (\mathbf{h}_1^H\mathbf{h}_2)^2/(\|\mathbf{h}_1\|^2\|\mathbf{h}_j\|)^2$ [17]. The numerical results are shown in Table I, where we also compute the mean value of ζ_i, $E(\zeta)$ and the rooted mean square deviation $\mathrm{RMSD}(\zeta) = \left(\sum_{i=1}^{N_{tot}}(\zeta_i - E(\zeta))/N_{tot}\right)^{0.5}$. We see that when $d_2 = 5$m and $d_2 = 10$m, not only there is high probability that ϵ_{ND} will happen, the performance gap between SCBF and ZFBF is also quite large (e.g for $d_2 = 5$m, ZFBF can use as much as 30% more transmission power than SCBF). This suggests that in the non-degraded case with channel asymmetry less prominent, SCBF does not outperform ZFBF all the time and there is a great amount of randomness w.r.t. which scheme would outperform the other one.

In [17], the proposed algorithm works in a low-complexity manner only for the degraded problem. In the presence of ϵ_{ND}, they have to solve the original problem (8) with a high computational burden. To obviate this problem, they propose to use ZFBF in the non-degraded case and therefore a performance loss is hard to be avoided. In our work, we compare the performance between ZFBF and SCBF every time when ϵ_{ND} occurs and choose the better one of them. The easy implementation of this performance comparison is made possible by the novel low-complexity algorithm, that solves (8) even in the presence of ϵ_{ND}, as we shall describe next.

2) Solving the Original Problem for the Non Degraded Case: In this part we focus on the case when (8c) is not satisfied by the optimal solution of $\mathbf{P}_D$. At the optimal point of $\mathbf{P}_D$, it can be written that

$$\text{(8b)} \Rightarrow K_{21\mathrm{od}}^2 = \frac{1}{\gamma_1}K_{1\mathrm{od}}^2 - \frac{\sigma_1^2}{s_0}, \tag{38a}$$

$$\text{(8c) disobeyed} \Rightarrow (K_{1\mathrm{od}} + K_{2\mathrm{od}})^2 < \gamma_1(\gamma_2\sigma_2^2 + \sigma_2^2), \tag{38b}$$

$$\text{(8d)} \Rightarrow K_{21\mathrm{od}} + K_{22\mathrm{od}} = \sqrt{\gamma_2\sigma_2^2}, \tag{38c}$$

where (38a), (38c) are due to the fact that (8b) and (8d) are active at the optimal point of $\mathbf{P}_D$. At the optimal point of the original problem of $\mathbf{P}_D$, the following equations are satisfied

$$\text{(8b)} \Rightarrow K_{21}^2 \le \frac{1}{\gamma_1}K_1^2 - \frac{\sigma_1^2}{s_0}, \tag{39a}$$

$$\text{(8c) does hold} \Rightarrow K_1 + K_2 = \gamma_1(\gamma_2\sigma_2^2 + \sigma_2^2), \tag{39b}$$

$$\text{(8d)} \Rightarrow K_{21} + K_{22} = \sqrt{\gamma_2\sigma_2^2}, \tag{39c}$$

where $s_0 = \frac{|r_{11}^* r_{12}|^2}{|r_{12}|^4}$. Our main goal is to find out how to satisfy (39a)-(39c) starting from (38a)-(38c). More specifically, K_1, K_2 should be varied appropriately to boost the left term value of inequality (38b). Another point worthwhile mentioning is that by comparing (38a)-(38c) and (39a)-(39c), we observe that there would be no benefit to make K_1 smaller than $K_{1\mathrm{od}}$ since $K_{1\mathrm{od}}$ is the optimal solution for $\mathbf{P}_D$. By further considering that $K_{2\mathrm{od}} = 0$, we propose to quantify the efficiency of increasing K_1 and K_2 starting from $K_{1\mathrm{od}}$ and $K_{2\mathrm{od}}$. One direct approach is by computing the partial derivative of the objective function, denoted as $f = \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2$, with respect to K_1 and K_2. $\partial f/\partial K_i$ could be viewed as the price to be paid if K_i were increased and the efficiency of the change is considered as high when the derivative is small. At the optimal point, it should be expected that $\partial f/\partial K_1 = \partial f/\partial K_2$.

For K_2, a simple result is available as it only influences (8c) and the variation of K_2 is actually independent of $\mathbf{w}_2$. Therefore from (10), we could write that

$$\frac{\partial f}{\partial K_2} = \frac{\partial\|\mathbf{w}_1\|^2}{\partial K_2} = \frac{2K_2}{|r_{22}|^2}, \quad K_2 \ge 0. \tag{40}$$

For K_1, more complex analysis is needed as increasing K_1 would not only affect $\mathbf{w}_1$, but at the same time due to (39a) increases K_{21} and therefore produces a possibility of a change in $\mathbf{w}_2$. From (39c), $\mathbf{w}_2$ is a quadratic function of K_{21}

$$rCl\|\mathbf{w}_2\|^2 = \left\{ \frac{K_{21}^2}{|r_{12}|^2} + \frac{K_{22}^2}{|r_{22}|^2} \middle| K_{21} + K_{22} = \sqrt{C} \right\} = g(K_{21}), \tag{41}$$

where $C = \gamma_2\sigma_2^2$. We denote the minimum point of g as $K_{21\mathrm{m}} = |r_{12}|^2\sqrt{C}/(|r_{12}|^2+|r_{22}|^2)$, which corresponds to the maximum ratio transmission (MRT) solution of $\mathbf{w}_2$. Notice that in the process of solving $\mathbf{P}_D$, due to the necessity of controlling the interference of $\mathbf{w}_2$ to User 1, $K_{21\mathrm{od}}$ should be less than $K_{21\mathrm{m}}$ (recall that $\mathbf{h}_1 = r_{11}\mathbf{v}_1$ and a larger K_{21} results in a larger correlation between $\mathbf{w}_2$ and $\mathbf{h}_1$). Therefore the growth of K_1 leads K_{21} to change towards its MRT solution. However, when K_{21} reaches $K_{21\mathrm{m}}$, (8b) or (39a) will no longer be activated and K_{21} then becomes independent of K_1. We express such a mathematical relationship as follows

$$\frac{\partial\|\mathbf{w}_2\|^2}{\partial K_1} = \frac{\partial g}{\partial K_{21}}\frac{\partial K_{21}}{\partial K_1}, K_{1\mathrm{od}} < K_1 \le K_{1\mathrm{m}} \tag{42}$$

$$\frac{\partial g}{\partial K_{21}} = 2\left(\frac{1}{|r_{12}|^2} + \frac{1}{|r_{22}|^2}\right)K_{21} - \frac{2\sqrt{C}}{|r_{22}|^2}, \tag{43}$$

$$\frac{\partial K_{21}}{\partial K_1} = \frac{1}{\gamma_1\sqrt{\frac{1}{\gamma_1} - \frac{\sigma_1^2}{s_0 K_1^2}}}, \tag{44}$$

where $s_0 = \frac{|r_{11}^* r_{12}|^2}{|r_{12}|^4}$, $K_{1\mathrm{m}}$ is related to $K_{2\mathrm{m}}$ by $K_{1\mathrm{m}}^2/\gamma_1 - \sigma_1^2/s_0 = K_{21\mathrm{m}}^2$, (43) is due to (41) and (44) is due to (38a). Notice that the term computed in (43) is negative and approaches zero as K_{21} tends towards $K_{21\mathrm{m}}$. We conclude that $\frac{\partial f}{\partial K_1}$ equals to

$$\frac{\partial\|\mathbf{w}_1\|^2}{\partial K_1} + \frac{\partial\|\mathbf{w}_2\|^2}{\partial K_1} = \begin{cases} 2\dfrac{K_1}{|r_{12}|^2} + \dfrac{\partial g}{\partial K_{21}}\dfrac{\partial K_{21}}{\partial K_1} & , K_{1\mathrm{od}} \le K_1 \le K_{1\mathrm{m}} \\ 2\dfrac{K_1}{|r_{12}|^2} & , K_1 > K_{1\mathrm{m}}. \end{cases} \tag{45}$$

The effect of increasing K_1 and K_2 is captured through (45) and (40). It is important to see that both of them are monotonically increasing. Note that by substituting K_{21} for K_1 in $\frac{\partial g}{\partial K_{21}}$, we can obtain $\left((\frac{1}{|r_{12}|^2} + \frac{1}{|r_{22}|^2})\sqrt{\frac{1}{\gamma_1}K_1^2 - \frac{\sigma_1^2}{s_0}} - \frac{\sqrt{C}}{|r_{22}|^2}\right)$. As can be derived from earlier discussions, this term should be negative when $K_1 < K_{1\mathrm{m}}$ and approaches zero as K_1 tends towards $K_{1\mathrm{m}}$, where $K_{1\mathrm{m}}$ is related to $K_{2\mathrm{m}}$ by $K_{1\mathrm{m}}^2/\gamma_1 - \sigma_1^2/s_0 = K_{21m}^2$. On the other hand, by checking (44), we see that $\partial K_{21}/\partial K_1$ is a positive decreasing term with respect to K_1. Therefore the product of the aforementioned two terms should be a monotonic increasing function of K_1 and $\partial f/\partial K_1$ should also be monotonic increasing in its own domain. We then argue that the optimal solution is achieved if the following condition is met when $K_1 \ge K_{1\mathrm{od}}, K_2 \ge K_{2\mathrm{od}}$

$$rCl\frac{\partial f}{\partial K_2} = \frac{\partial f}{\partial K_1}, \tag{46}$$

$$K_1 + K_2 = \sigma_2\sqrt{\gamma_1(\gamma_2+1)} = s_1. \tag{47}$$

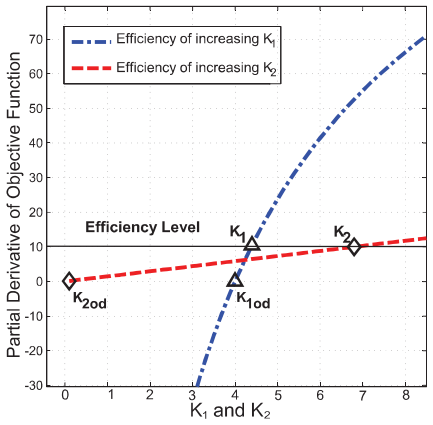


Fig. 2. Efficiency comparison between the optimization variables.

Suppose (46) does not hold. Through the definition of partial derivative, better solution could be found by increasing the more efficient optimization variable, while decreasing the other one. Equation (47) follows from (8d). We show one typical case in Fig. 2 where the efficiency level is used to represent the condition of (46). It is interesting to note that satisfying (46) and (47) at the same time is analogous to procedures implemented in the water-filling algorithm, where the allocation principle for the transmission power is also related to the formation of a virtual level. Problem (8) now boils down to finding K_1 and K_2 that satisfy (46), (47), which is equivalent to solving the following equation under the condition that $K_1 \ge K_{1\mathrm{od}}$,

$$\begin{aligned} h(K_1) = {} & K_1\left(\frac{1}{|r_{12}|^2} + \frac{1}{|r_{22}|^2}\right) \\ & + I_+(K_1 - \mathrm{K}_{1\mathrm{m}})\left(\frac{1}{\gamma_1\sqrt{\frac{1}{\gamma_1} - \frac{\sigma_1^2}{\mathrm{s}_0\mathrm{K}_1^2}}}\right) \\ & \times \left((\frac{1}{|r_{12}|^2} + \frac{1}{|r_{22}|^2})\sqrt{\frac{1}{\gamma_1}K_1^2 - \frac{\sigma_1^2}{s_0}} - \frac{\sqrt{C}}{|r_{22}|^2}\right) \\ & - \frac{s_1}{|r_{22}|^2} = 0, \end{aligned} \tag{48}$$

where I_+ is indicator function for nonnegative real numbers,

$$I_+(u) = \begin{cases} 0, & \text{if } u \ge 0 \\ 1, & \text{if } u < 0. \end{cases} \tag{49}$$

It can be checked that $h(K_1)$ is a piecewise smooth, continuous and most importantly, monotonic increasing function of K_1. We conclude the above process in Algorithm 2. A detailed analysis of the searching for the solution of equation (48) is provided in the next section.

D. Computational Complexity Analysis

Compared to the second stage in Algorithm 1, which requires multiple computations of a matrix inverse, the main task in Algorithm 2 is to solve the non-linear one dimensional equation represented by (48). Due to the monotonicity and

Algorithm 2 Incremental Optimization Algorithm

1: Solve $\mathbf{P}_D$ by setting $\lambda_2 = 0$ in Stage-A of Algorithm 1.
2: Set $K_{1\mathrm{od}}$ as the initial point and search for the solution of equation (48) by using Newton's method [22]. The final solutions for $\mathbf{w}_i^{\mathrm{SCBF}}, i = 1, 2$ can be obtained by combining equations (39a), (39c) and (47).
3: Compute $\mathbf{w}_i^{\mathrm{ZFBF}}, i = 1, 2$ according to (37).
4: Choose between $\mathbf{w}_i^{\mathrm{ZFBF}}, i = 1, 2$ and $\mathbf{w}_i^{\mathrm{SCBF}}, i = 1, 2$ by comparing the resulted transmission power.

continuity of $h(K_1)$, we can simplify the problem by considering whether $K_1 \geq K_{1\mathrm{m}}$ is satisfied. We can check this condition by simply solving the following one-dimensional linear equation of K_1,

$$K_1 \left(\frac{1}{|r_{12}|^2} + \frac{1}{|r_{22}|^2} \right) - \frac{s_1}{|r_{22}|^2} = 0, \tag{50}$$

and if the solution $K_1 > K_{1\mathrm{m}}$, then it is the right solution. Otherwise, we need to solve the nonlinear equation $h(K_1) = 0$ for $K_{1od} \leq K_1 \leq K_{1\mathrm{m}}$, on which $-h(K_1)$ has positive and continuous second order derivatives and is therefore convex. By applying Newton's method [22], the root to $h(K_1) = 0$ can be computed efficiently. The global convergence is guaranteed by the convexity of $-h(K_1)$ and the continuity of its gradient on $K_{1\mathrm{od}} \leq K_1 \leq K_{1m}$, as proved in [22, Th. 13.3.7].

To fully characterize the computational complexity of Algorithm 2, we need to also analyze the implementation of Stage-A in Algorithm 1. Note that the matrix inverse required in Stage-A of Algorithm 1 can actually be greatly simplified by using Sherman-Morrison formula [23]

$$\mathbf{G}^{-1} = (\mathbf{I} + \lambda_i \mathbf{h}_i \mathbf{h}_i^H)^{-1} = (\mathbf{I} - \frac{\lambda_i \mathbf{h}_i \mathbf{h}_i^H}{1 + \mathbf{h}_i^H \mathbf{h}_i}). \tag{51}$$

Therefore, the matrix inverse can be computed with complexity $\mathcal{O}(N)$, where N is the number of antenna. The computational complexity for the searching stage, however, does not grow with N and can be considered as constant in an asymptotic analysis.

By transforming problem (8) into the standard form of SOCP (as shown in Appendix B), the worst-case complexity for the SOCP and SDP methods can be obtained. We conclude the complexity results in Table II, where m is the number of constraints in (B.6a), n_i is the number of rows in A_i plus one [8]. As shown in (A.3), $m = 4, n_1 = n_2 = n_4 = 3$, $n_3 = 2$. Since both the SDP and SOCP methods require a complexity that grows quadratically with the number of antennas, it can be expected that Algorithm 2 has much lower computational complexity when the number of antennas is large. Table II also indicates that the SOCP method has lower worst-case complexity than its SDP counterpart.

TABLE II
COMPUTATIONAL COMPLEXITY COMPARISON

Candidate Schemes	Worst-case Complexity
SOCP	$O(\sqrt{m}) \cdot O(N^2 \sum_{i=1}^4 n_i)$
SDP	$O(\sqrt{\sum_{i=1}^4 n_i}) \cdot O(N^2 \sum_{i=1}^4 n_i^2)$
Algorithm 2	$O(N)$

IV. Extension to More Than Two Receivers and to Rate Region Optimization

In this section, we first generalize the technique in Section III to cases where more than two layers of information are overlaid. In general cases, problem (6) is hard to be solved optimally. However, for $M = 3$ we do provide efficient algorithms that approach the optimal performance with reasonable computational complexity. We also generalize our method to the computation of the two user rate region achieved by SCBF. Finally to deal with larger networks, we design a hybrid precoding scheme that combines ZFBF and SCBF.

A. Sub-Optimal Multi-Layer Beamforming With Superposition Coding

First, we show that for each optimization problem in the form of (52), a set of coefficients can be found to transform the original problem into the SOCP without loss of optimality.

Proposition 3: Consider the following non-convex program

$$\min_{\mathbf{x}} \ \|\mathbf{x}\| \ \ \text{s.t.} \ \|\mathbf{A}_i\mathbf{x}\|^2 \leq |\mathbf{a}_i\mathbf{x}|^2, \ i = 1, \ldots, L, \tag{52}$$

where $\mathbf{A}_i\mathbf{x}$ and $\mathbf{a}_i\mathbf{x}$ form affine functions of $\mathbf{x}$, the objective function is the ℓ_2-norm of $\mathbf{x}$. We could always find a set of rotation coefficients of fixed values, $e^{j\phi_i}, i = 1, \ldots, L$, through which a new convex program, as shown in (53)*, is constructed to achieve the same optimal objective value as that achieved by the original one, i.e.* (52)*.*

$$\min_{\mathbf{x}} \ \|\mathbf{x}\| \ \ \text{s.t.} \ \|\mathbf{A}_i\mathbf{x}\| \leq \Re(e^{j\phi_i}\mathbf{a}_i\mathbf{x}), \quad i = 1, \ldots, L. \tag{53}$$

Proof: We first prove that the transformed problem will achieve the optimal value of the original one. Suppose $\mathbf{x}_{\mathrm{o}}$ is the optimal solution of (52). Properly rotate each term on the right side of the constraint in (52), i.e. $\mathbf{a}_i\mathbf{x}$, and we could always have

$$C\Re(e^{j\theta_i}\mathbf{a}_i\mathbf{x}_o) \geq \|\mathbf{A}_i\mathbf{x}\|, \quad i = 1, \ldots, L. \tag{54}$$

Note that this means that $\mathbf{x}_{\mathrm{o}}$ should lie in the feasible set of problem (53) with the same set of rotation coefficients as in (54). We finish our proof by showing that the feasible set of (53) is included by the feasible set of the original problem. For every $\mathbf{x}$ providing a feasible solution to (53), it is easy to see that it also belongs to the feasible set of problem (52). □

Note that problem (8) is exactly of the same form as (52). Hence to solve problem (8), we first find a set of rotation coefficients and then transform the original problem. However, in scenarios where more than two users are involved, the rotation coefficients are difficult to be predetermined. We first write the QR decomposition $\mathbf{H} = \mathbf{V}\mathbf{R} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M)\mathbf{R}$. To guarantee the existence of such a QR decomposition and the invertibility of matrix $\mathbf{R}$, we need to assume here that $N \geq M$, i.e the number of transmit antennas is not smaller than the number of users. The beamforming vectors can be expressed

as $\mathbf{w}_i = \mathbf{V}\boldsymbol{\alpha}_i = \alpha_{i1}\mathbf{v}_1 + \ldots + \alpha_{iM}\mathbf{v}_M$. Corresponding to the right side terms of the constraints in (53), we can write

$$\boldsymbol{\Lambda}_i\mathbf{H}^H\mathbf{w}_i = \boldsymbol{\Lambda}_i\mathbf{R}^H\boldsymbol{\alpha}_i = \mathbf{p}_i, \quad i = 1, \ldots, M, \tag{55}$$

where $\boldsymbol{\Lambda}_i = \mathrm{Diag}(e^{j\phi_1}, \ldots, e^{j\phi_M})$ is constituted by the rotation coefficients. Then the beamforming vector $\mathbf{w}_i$ can be rewritten as $\mathbf{w}_i = (\boldsymbol{\Lambda}_i\mathbf{R}^H)^{-1}\mathbf{p}_i = \mathbf{A}\boldsymbol{\Lambda}_i'\mathbf{p}_i$, where $\mathbf{A} = (\mathbf{R}^H)^{-1}$, $\mathbf{p}_i$ is non-negative real and $\boldsymbol{\Lambda}' = \boldsymbol{\Lambda}^H$. We may now interpret each rotation coefficient $e^{j\phi_i}$ as a factor that controls the common phase shifter added to the component of $\mathbf{w}_i$ along the i^{th} column of $\mathbf{A}$, while $\mathbf{p}_i$ controls $\mathbf{w}_i$'s magnitude.

Fixing $\boldsymbol{\Lambda}_i$, there exists an SOCP transformation of the original minimum power optimization with QoS constraints, as expressed in the below problem,

$$OBJ = \min_{\mathbf{p}_1,\cdots,\mathbf{p}_M} \sum_{i=1}^{M} \mathbf{p}_i^H \boldsymbol{\Lambda}_i^{'H}\mathbf{A}^H\mathbf{A}\boldsymbol{\Lambda}_i'\mathbf{p}_i \tag{56a}$$

$$\text{s.t. } p_{ji} \geq \sqrt{\gamma_j \sum_{k\geq j}^{M} p_{ki}^2}, \quad j \in [1, M-1],\ i \in [1, M], \tag{56b}$$

$$p_{MM} \geq \sqrt{\gamma_M}, \tag{56c}$$

where p_{ji} is the i^{th} element of vector $p_j, j = 1, 2$. Problem (56) is clearly of an SOCP form, where we are only searching over $\mathbf{p}_i$. In general the optimality is not preserved in the transformed problem since the feasible set of (56) is only included by that of its original problem. However, according to Proposition 3, with the right choice of $\boldsymbol{\Lambda}'$, (56) has the same optimal value as the original problem (6). This is just what has happened to the two-user case in problem (8), where we predetermine the optimal values of $\boldsymbol{\Lambda}_i$, which can be explicitly written as $\boldsymbol{\Lambda}_1 = \mathrm{Diag}(e^{j\phi_{12}}, 1), \boldsymbol{\Lambda}_2 = \mathrm{Diag}(1, 1)$, and are used to transform (8) into (12) without loss of optimality (remember that ϕ_{12} is the phase of $\mathbf{h}_1^H\mathbf{h}_2$).

When more than two users are considered, the optimal matrix $\boldsymbol{\Lambda}'$ becomes difficult to be predetermined. To illustrate this point, first compute the total transmission power as

$$rCl\mathbf{w}_i^H\mathbf{w}_i = \|(\boldsymbol{\Lambda}_i\mathbf{R}^H)^{-1}\mathbf{p}_i\|^2 = \mathbf{p}_i^H\boldsymbol{\Lambda}_i^{'H}\mathbf{A}^H\mathbf{A}\boldsymbol{\Lambda}_i'\mathbf{p}_i \triangleq f, \tag{57}$$

Now, consider the optimization problem of (58),

$$\min_{\boldsymbol{\Lambda}_i', \mathbf{p}_i = p_{ic}} \quad f(\mathbf{p}_i, \boldsymbol{\Lambda}_i'), \tag{58}$$

where the function f is minimized over $\boldsymbol{\Lambda}'$ with each $\mathbf{p}_i$ fixed to be certain positive number p_{ic}. When the optimal solution of (58) is independent of the choice of p_{ic}, the rotational coefficients can be predetermined. It turns out that this is only the case for $M = 2$. As shown by (59) that computed the objective function when $M = 2$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the columns of $\mathbf{A}$,

$$\|e^{-j\phi_1}p_{i1}\mathbf{a}_1 + e^{-j\phi_2}p_{i2}\mathbf{a}_2\|^2 = p_{i1}^2\mathbf{a}_1^H\mathbf{a}_1 + p_{i2}^2\mathbf{a}_2^H\mathbf{a}_2 + p_{i1}p_{i2}\mathbf{a}_1^H\mathbf{a}_2e^{j(\phi_1-\phi_2)} + p_{i1}p_{i2}\mathbf{a}_2^H\mathbf{a}_1e^{j(\phi_2-\phi_1)}, \tag{59}$$

it can be seen that the global minimum of (59) is achieved when $\phi_1 - \phi_2 = \psi_{12} + \pi$, where we use ψ_{ij} to represent the phase of $\mathbf{a}_i^H\mathbf{a}_j$. This is so because this special choice of $\phi_1 - \phi_2$ makes the last two terms of (59), which we defined as the cross terms, become non-positive reals (for example, $p_{i1}p_{i2}\mathbf{a}_1^H\mathbf{a}_2e^{j(\phi_1-\phi_2)} = p_{i1}p_{i2}|\mathbf{a}_1^H\mathbf{a}_2|e^{j\pi} = -p_{i1}p_{i2}|\mathbf{a}_1^H\mathbf{a}_2|$). For $M = 3$, the objective function becomes $\|e^{-j\phi_1}p_{i1}\mathbf{a}_1 + e^{-j\phi_2}p_{i2}\mathbf{a}_2 + e^{-j\phi_3}p_{i3}\mathbf{a}_3\|^2$. Different from the two-user case, the optimal rotational coefficients are now not always independent of the positive number p_{ij}. Therefore for $M \geq 3$, we propose a sub-optimal algorithm to solve problem (6) by iteratively optimizing over $\boldsymbol{\Lambda}_i'$ and $\mathbf{p}$. Note that $\mathbf{w}_i$ could be obtained through $\mathbf{w}_i = (\boldsymbol{\Lambda}_i\mathbf{R}^H)^{-1}\mathbf{p}_i$, after the implementation of Algorithm 3. In Appendix VI, we show that for $M = 3$, problem (58) boils down to an one-dimensional searching over the finite region $[0, 2\pi]$. As M increases, the complexity grows exponentially. The convergence proof is as follows.

Algorithm 3 Minimum Power SCBF With Multiple Layers

Initialization: randomly generate $\boldsymbol{\Lambda}_i^{'0}$, $i = 1, \cdots, M$, t:=0

Repeat until: convergence or required number of iterations

1: Solve problem (56). The optimal solution is denoted as $\mathbf{p}_{io}$, $i = 1, \cdots, M$.

2: Solve problem (58) for $i = 1, \cdots, M$. The optimal solutions are denoted as $\boldsymbol{\Lambda}_{io}^{'t}, i = 1, \cdots, M$ respectively.

3: Set $\boldsymbol{\Lambda}_i^{'t+1} = \boldsymbol{\Lambda}_{io}^{'t}, i = 1, \cdots, M$. $t = t + 1$.

Proof: First denote the objective value achieved by solving (56) in the i_{th} iteration as OBJ_i. Due to (58), it could be deduced that $\mathrm{OBJ}_i \geq \mathrm{OBJ}_{i+1}$. It could also be easily seen that OBJ_i is lower bounded by zero. Therefore we could conclude by the monotonicity and boundedness of OBJ_i that the iterations will converge to a stationary point. □

B. Optimal Results for 3-Layer SCBF and Two-User Rate Region Computation

We first show that by extending the previous trick for the 2-user case and under Assumption 1, optimal solutions to problem (6) with $M = 3$ can be found with relatively low complexity.

Assumption 1: The QoS constraint $SINR_3^1 \geq \gamma_1$ is inactive for the optimal solution to problem (6) *with* $M = 3$.

Assumption 1 indicates that the strongest user could so easily decode the upper layer information to the extent that the QoS constraint $\mathrm{SINR}_3^1 \geq \gamma_1$ can be ignored. Such an assumption is defined as cross-layer degradedness (CLD). We then transform the remaining constraints into the SOC types just as we did in (12) with one exception

$$ll\Re(e^{j\theta}\mathbf{h}_2^H\mathbf{w}_2) \geq \gamma_2'\sqrt{|\mathbf{h}_2^H\mathbf{w}_3|^2 + \sigma_2^2}. \tag{60}$$

where $\gamma_2' = \sqrt{\gamma_2}$. The added coefficient $e^{j\theta}$ in the left term of (60) has an undetermined phase. We argue that under CLD, the optimal solution to (6) can be found by solving the newly formed SOCP with a correct choice of the phase value θ. Motivated by (11), instead of traversing multiple values of θ, we propose to solve only the SOCP with θ chosen as the phase of $\mathbf{h}_2^H\mathbf{h}_3$. We denote such a heuristic method as SOCP under CLD. When CLD does not hold, we simply add to $\mathbf{w}_1$

the signal component that is orthogonal to $\mathbf{h}_1, \mathbf{h}_2$, so that the QoS requirement is satisfied. Later through simulations, we demonstrate that with practical parameter setups the performance of our proposed method stays very close to the optimal one.

For the rate region computation problem (7), we can similarly transform the non-convex constraints into SOC constraints, as we deal with problem (6). It would also be possible to design more dedicated algorithms like the process we perform in Section III-B and Section III-C, which will be interesting future work.

C. Multi-User Hybrid Beamforming With User Grouping

In practice, far greater networks than the two user case are usually met. To bound the computational complexity, it would be attractive to combine SCBF with orthogonal multiplexing methods. One such combination is through user grouping. While SCBF is implemented within each group, zero-forcing is used to suppress the interference between groups. We partition the total users into $M_g = M/g$ mutually exclusive sub-groups of size g, which indicate the number of information layer multiplexed in the power domain. The i^{th} group is represented by the ordered set $\mathcal{S}_i = \{u_{i_1}, \ldots, u_{i_g}\}$, the ordering of which is used to determine the decoding sequence. The scalar g represents the number of users multiplexed in power domain. Notice that $\mathcal{S} = \cup_i^{M_g} \mathcal{S}_i$, and $\forall i \neq j, \mathcal{S}_i \cup \mathcal{S}_j = \phi$, where ϕ denotes null set. Let $\mathbf{h}_{i_j}, \mathbf{w}_{i_j}$ denote the channel vector and the precoding vector associated with u_{i_j} respectively. The columns of $\mathbf{G}[i]$, which is a sub-matrix of $\mathbf{H}$, correspond to the channel vectors associated with users in $\mathcal{S}_i$. We use $\mathbf{T}[i] \in \mathbb{C}^{N \times (M-g)}$ to denote the remaining part of $\mathbf{H}$ after removing $\mathbf{G}[i]$ from it. In order to remove the inter-group interference, we need to ensure that $\mathbf{w}_{i_j} \in \mathcal{T}[i]^{\perp}$.

We then derive the expressions for the effective channels. We introduce matrix $\mathbf{V}[i]$, which forms an orthonormal basis of $\mathcal{T}[i]^{\perp}$. To ascertain $\mathbf{V}[i]$, write the QR decomposition as $\mathbf{VR} = \mathbf{F} = \mathbf{PG}$, where $\mathbf{P}$, the projection matrix is obtained as

$$\mathbf{P} = \mathbf{I} - \mathbf{T}(\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H. \tag{61}$$

Note that we omit the corresponding sub-scripts in the above equations for convenience. From $\mathbf{T}^H\mathbf{VR} = \mathbf{T}^H\mathbf{PG} = \mathbf{0}$, it can be seen that $\mathbf{T}^H\mathbf{V} = \mathbf{0}$. Therefore Col($\mathbf{V}$) is orthogonal to $\mathbf{T}$. By using the Hermitian nature of $\mathbf{P}$, the following equations are further obtained,

$$\mathbf{R} = \mathbf{V}^H\mathbf{F} = \mathbf{V}^H\mathbf{PG} = \mathbf{PV}^H\mathbf{G} = \mathbf{V}^H\mathbf{G}, \tag{62}$$

where the last equation is due to the fact that $\mathbf{P}$ and $\mathbf{V}$ represent the same linear sub-space. Now let $\mathbf{w} = \mathbf{V}\hat{\mathbf{w}}$ and we write

$$\mathbf{G}^H\mathbf{w} = \mathbf{G}^H\mathbf{V}\hat{\mathbf{w}} = \mathbf{R}^H\hat{\mathbf{w}}, \tag{63}$$

from which we could conclude that $\mathbf{R}$ is the effective channel gains with respect to $\hat{\mathbf{w}}$. By resorting to the techniques proposed in Section III or Section IV, the corresponding minimum power optimization problem could be solved.

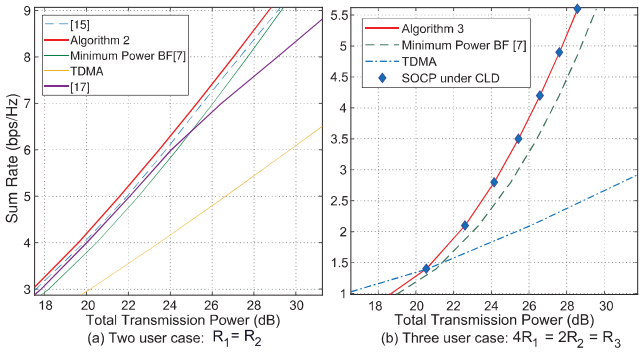


Fig. 3. Sum rate versus total transmission power for the two and three user case in Rayleigh fading channel with $N = 5, d_3 = 1m, d_2 = 10m, d_1 = 20m$.

Note that the precoding vector $\mathbf{w}$ is coordinate transformed version of $\hat{\mathbf{w}}$. Let $\mathcal{T}[i]$ contain the set of users whose sub-channels are orthogonal to $\mathbf{V}[i]$. If $\mathcal{T}[i]$ is chosen as the set of users that correspond to the remaining part of $\mathbf{H}$ after removing $\mathbf{G}[i]$ from it, then any transmitted signal within the column space of $\mathbf{V}[i]$ should not interfere any user from the other groups. Applying this choice of $\mathcal{T}[i]$ for all members in group $\mathcal{S}_i$, results in a complete elimination of the inter-group interference. We denote such a scheme as Hybrid BF.

D. User Grouping Strategy

Since the effective channel $\mathbf{R}[i]$ is the projection of $\mathbf{G}[i]$ onto $\mathcal{T}[i]^{\perp}$, having users whose channels are highly correlated in the different groups would lead to a large performance loss for ZFBF. To bound the complexity, we use a simple greedy grouping method. First according to the large-scale attenuations, or equivalently the distances, users are classified into g layers $L_1, \cdots, L_g$. Without loss of generality, suppose $g = 2$. Randomly pick one user from the first layer L_1, whose channel gain is denoted as $\mathbf{h}_i$. Then from those users in L_2, choose the one with $\mathbf{h}_j$ that yields the largest value of $|\mathbf{h}_i^H\mathbf{h}_j|/(\|\mathbf{h}_i\|\|\mathbf{h}_j\|)$. Place these two users in the same sub-group and move on to implement the same procedure for the remaining users.

V. Numerical Simulation and Analysis

In this section, we implement numerical stimulations to demonstrate the superiority of SCBF over the orthogonal transmission methods in both Rayleigh and Rician fading channels. Throughput performance achieved by SCBF for both the two-user and the multi-user cases is investigated. Finally, we compute the two-user rate region to further characterize its performance trend.

A. Numerical Results of SCBF in Rayleigh Fading Channel

Focusing on Rayleigh fading, the results have been averaged over the performance achieved by 100 independent channel realizations. Results concerning the relationship between the total transmission power and the sum rate are shown in Fig. 3(a) for the two-user case. It can be observed that Algorithm 2 in general achieves the best performance, with a power gain up to 1dB for the high SNR region over the [7].

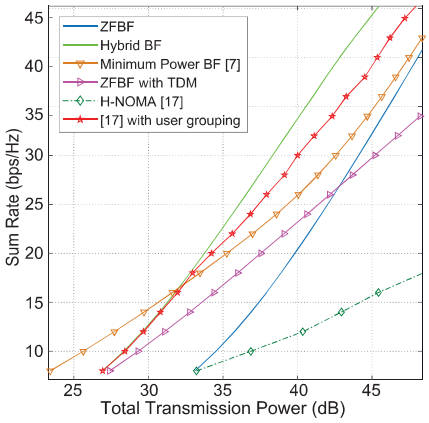


Fig. 4. Comparisons between different beamforming schemes in multi-user Rayleigh channels.

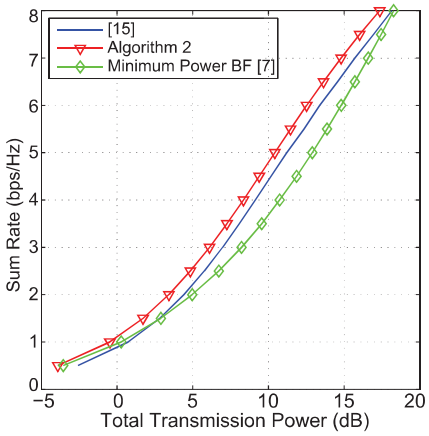


Fig. 5. Rician fading channels with $v_1 = 0.1, v_2 = 100$.

The heuristic algorithm in [15] and [17] are strictly outperformed by our method. Both the performance of Algorithm 3 with $M = 3$ and that achieved by the method proposed in Section III.B are shown in Fig. 3(b). The number of iterations for Algorithm 3 is fixed to be 5. The two methods yield very similar performance, both outperforming the other two orthogonal schemes at higher rates. Meanwhile we find Assumption 1 is satisfied over 90% of the time and when satisfied the performance gap between SOCP under CLD and the optimal solution is negligible. At lower rates, since the noise becomes the dominated factor affecting the transmission rate, TDMA achieves the best performance.

Among the five candidate schemes that we are comparing in our simulations, except for TDMA and [17], all the other three methods employ certain types of iterative algorithms. The method applied in [15] is based on the nonlinear Gauss-Seidel (GS) algorithm. At each iteration, an unconstrained optimization that further demands iterative searching is needed. Compared to this, both Algorithm 2, and the conventional optimization in [7] require solving fixed point equations. Since one more layer of iterations are needed in [15], the running time required by [15] is generally higher than Algorithm 2 and [7]. For the two user case, [7] needs to solve two fixed point equations while Algorithm 2 only requires solving one. Therefore, our method can be considered as more computational efficient in the sense of having fewer equations to solve.

B. Multi-User Hybrid Beamforming With User Grouping

In this section, two groups of users are considered with one distributed around 1 meter away from the BS and the other around $d = 5$ meters away from the BS. We assume an even rate allocation and that the number of transmitter antennas and users are both 8. Setting $g = 2$, the users are partitioned into $M/g = 4$ sub-groups. The user grouping strategy will be used as Section IV-D. We observe in Fig. 4(a) that Hybrid BF achieves significant performance gain over the other schemes when the sum rate is over 16 bit/s/Hz, while [7] achieves better results at lower rates. We also test the performance of the combination of Zero-forcing and TDM, i.e. ZFBF is applied to control the inter-group interference while TDM is used within a group. This scheme yields similar performance as Hybrid BF at low SNRs. The dotted line is H-NOMA proposed in [17] while the starred line combines the sub-optimal NOMA beamforming in [17] with user grouping. It is clearly seen that the performance gain of Hybrid BF comes partially from applying optimal solution and provides extra gain from user grouping.

C. SCBF in Rician Fading Channels

We assume d/λ to be 0.5 and a even rate allocation ($\gamma_1 = \gamma_2$). As seen in Fig. 5, SCBF outperforms the other two schemes at all tested rates. Note that we set $\delta\theta = \theta_2 - \theta_1 = \pi/6$. As $\delta\theta$ becomes larger, the two channels would become less correlated and the performance gap between the orthogonal BF scheme and SCBF is expected to decrease. It can be seen that in Rician fading, the channel degradedness of could be created not only by the difference in distances between the transmitter and the receivers, but also by appropriate geometric positions of users and their visibility to the transmitter.

D. Two User Rate Region of Beamforming With Superposition Codin

In this section, through numerical computations of the rate region achieved by SCBF, we quantify the impact on the efficiency of SCBF brought by transmission power, channel correlation and channel asymmetry. To facilitate the analysis, we adopt a special channel model

$$\mathbf{h}_i = \alpha_i \left(e^{j\frac{2\pi 0\Delta_i}{N}}, e^{j\frac{2\pi 1\Delta_i}{N}}, \ldots, e^{j\frac{2\pi(N-1)\Delta_i}{N}} \right) / N^{0.5}, \quad i = 1, 2, \tag{64}$$

where α_i and Δ_i are respectively the attenuation coefficient and the phase offset coefficient associated with the i^{th} user. Note that as the difference between Δ_1 and Δ_2, denoted as δ, changes from 0 to 1, $|\mathbf{h}_1^H \mathbf{h}_2|$ varies from 1 to 0. We fix the stronger user's channel gain α_2 to be 1. From Fig. 6, it can be observed that the rate regions achieved by the two user SCBF are relatively unchanged by the variation of the correlation between the channel gains. This explains why

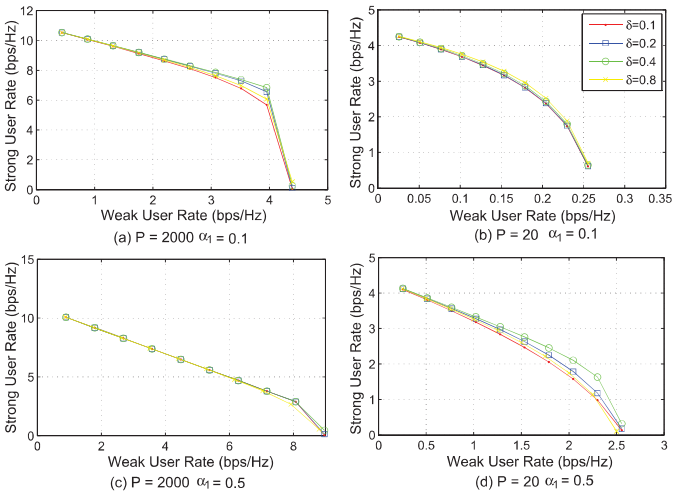


Fig. 6. Robustness of two-user SCBF against channel correlations.

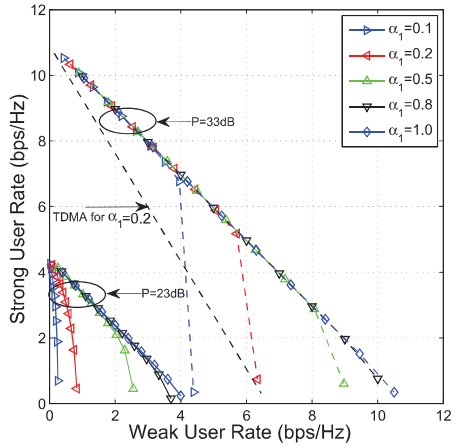


Fig. 7. Two user rate region achieved by beamforming with superposition coding.

SCBF provides better performance than orthogonal beamforming scheme since the latter one would degrade significantly as the channel correlation grows.

Finally by fixing δ, we show how the rate region is related to the transmission power and the degree of channel asymmetry (represented by different values of α_1). We simulate cases with total transmission powers of $P = 33$ dB (the dashed lines) and 23 dB (the solid lines) respectively and also with multiple levels of attenuation of the weaker channel. The comparison between SCBF and TDMA becomes very straightforward in Fig. 7, as the rate region achieved by TDMA is approximately the line segment between the first and the last point belonging to the rate region achieved by SCBF. For example, as can be seen in Fig. 7, when the channels are symmetric, i.e. $\alpha = 1$, the achievable region achieved by TDMA will coincide with that achieved by SCBF. More interestingly, it can be seen that the beginning part of a typical rate region consists of a unit-slope linear part, which means that the rate trade-off between the two users follows the case where TDMA is applied to symmetric channels. As the weaker user rate continues to increase, there is then a rapid deterioration of the strong user's rate.

VI. Conclusion

In this paper, we study the optimization algorithms for beamforming with superposition coding. A novel convex approximation method is proposed first, based on which we design several efficient algorithms to solve the power minimization and the rate maximization problems. We also propose to combine SCBF with ZFBF through user grouping. Optimal results are guaranteed to be found for the two user case and we experimentally show that a nearly optimal performance can be achieved for the three user case with practical problem setups. Numerical simulations demonstrate that the proposed methods outperform the baseline references significantly. Particularly in Rician fading channels, the SCBF transmission scheme could achieve a 2 to 3 dB power gain over the orthogonal multiplexing method at both moderate and high SNRs.

Appendix A
Proof of Proposition 2

To analyze the iterative process, i.e. Stage-B in Algorithm 1, we first rewrite the term $\frac{|\mathbf{h}_1^H\mathbf{G}^{-1}\mathbf{h}_2|^2}{|\mathbf{h}_2^H\mathbf{G}^{-1}\mathbf{h}_2|^2}$ in (30) by using Sherman-Morrison formula [23]. Define $\left(\mathbf{I}+\lambda_1\mathbf{h}_1\mathbf{h}_1^H\right)$ as $\mathbf{G}_1$, the inverse of which can be written as

$$\mathbf{G}_1^{-1} = (\mathbf{I}+\lambda_1\mathbf{h}_1\mathbf{h}_1^H)^{-1} = \mathbf{I} - \frac{\lambda_1\mathbf{h}_1\mathbf{h}_1^H}{1+\lambda_1\mathbf{h}_1\mathbf{h}_1^H}. \quad \text{(A.1)}$$

Again, by Sherman-Morrison formula, we write

$$l\mathbf{G}^{-1} = (\mathbf{G}_1+\lambda_2\mathbf{h}_2\mathbf{h}_2^H)^{-1} = \mathbf{G}_1^{-1} - \frac{\lambda_2\mathbf{G}_1^{-1}\mathbf{h}_2\mathbf{h}_2^H\mathbf{G}_1^{-1}}{1+\lambda_2\mathbf{h}_2^H\mathbf{G}_1^{-1}\mathbf{h}_2}. \quad \text{(A.2)}$$

Define $D_1 = \mathbf{h}_2^H\mathbf{G}_1^{-1}\mathbf{h}_2$ and $D_2 = \mathbf{h}_1^H\mathbf{G}_1^{-1}\mathbf{h}_2$. From (A.2), it can be shown that

$$l\mathbf{h}_i^H\mathbf{G}^{-1}\mathbf{h}_2 = D_1\left(1-\frac{D_1\lambda_2}{1+D_1\lambda_2}\right), \quad i=1,2. \quad \text{(A.3)}$$

Note that D_1 and D_2 are scalar functions of λ_1. Since $\mathbf{G}_1^{-1}$ is positive definite, D_1 should be real and positive. By multiplying $\mathbf{h}_1$ with an appropriate coefficient, D_2 can also be kept real. After substituting (A.3) into (30), we rewrite (30) as

$$\lambda_2 = u(\lambda_1) = \frac{1}{c_3}\left[(\gamma_1\gamma_2\left(\frac{D_2}{D_1}\right)^2+\gamma_1)^{0.5} - R_c\frac{a_{11}}{a_{12}}\right], \quad \text{(A.4)}$$

where we define the new function u to indicate that the right hand side of (30) is in fact a function only of λ_1 (not λ_2) since $D_2/D_1 = a_{12}/[a_{22}+(a_{11}a_{22}-a_{12}^2)\lambda_1]$. Notice that function u is a monotonically increasing function of $\lambda_1 > 0$. Then from (31), (32), λ_1 equals to

$$v(\lambda_2) = \frac{\gamma_1}{a_{11}a_{22}-a_{12}^2}\left(a_{22}+\gamma_1 a_{12}^2\frac{1}{(a_{11}a_{22}-a_{12}^2)\lambda_2 - a_{11}\gamma_1}\right).$$

Now after replacing λ_1 in (A.4) with the above function, i.e. $\lambda_2 = u(v(\lambda_2))$, we derive a fixed point iteration that is equivalent to the original iteration process in Algorithm 1.

Denoting t as the iteration index and $\mathcal{D}$ as the domain of $u(v(\lambda_2)$, the iteration can be interpreted as the difference equation, $\lambda_2^{t+1} = u(v(\lambda_2^t)) = t(\lambda_2^t), \lambda_2 \in \mathcal{D}$. Define $\mathcal{N}$ as a closed interval, such that $\mathcal{N} \subset \mathcal{D}$ and $\forall x \in \mathcal{N}, |t'(x)| < 1$. According to [24, Corollary 1.14], if $\lambda_2^0, \lambda_2^f \in \mathcal{N}$, the iteration must converge to the fixed point λ_2^f.

APPENDIX B
STANDARD SOCP FORM OF PROBLEM (8)

To characterize the worst-case complexity for the SOCP and SDP approaches for problem (8), transform the original problem into the following standard form of SOCP [8],

$$\min_{\mathbf{x}} \ \mathbf{f}^H\mathbf{x} \ \text{ s.t. } \ \|\mathbf{A}_i\mathbf{x}+\mathbf{b}_i\| \le \mathbf{c}_i^H\mathbf{x}, \quad i=1,\cdots,4, \quad \text{(B.6a)}$$

where $\mathbf{x} = [\mathbf{w}_1^H, \mathbf{w}_2^H, P]^H$, with P being the newly introduced real scalar variable, is the optimization variant, $f = [0,\cdots,0,1] \in \mathcal{C}^N$,

$$\mathbf{A}_1 = \begin{pmatrix} 0 & \cdots & \mathbf{h}_1^H & 0 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 0 & \cdots & \mathbf{h}_2^H & 0 \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

$$\mathbf{A}_3 = \mathbf{0}, \quad \mathbf{A}_4 = \begin{pmatrix} 1 & \cdots & 1 & 0 \\ 1 & \cdots & 1 & 0 \end{pmatrix},$$

$\mathbf{b}_1 = [0,\sigma_1]^H, \mathbf{b}_2 = [0,\sigma_2]^H], \mathbf{b}_3 = \sigma_2, \mathbf{b}_4 = [0,0]^H, \mathbf{c}_1 = \mathbf{c}_2 = 1/\sqrt{\gamma_1}[\mathbf{h}_1^H, 0,\cdots,0]^H, \mathbf{c}_3 = 1/\sqrt{\gamma_2}[0,\cdots,0,\mathbf{h}_2^H]^H, \mathbf{c}_4 = [0,\cdots,0,1]^H$.

APPENDIX C
SOLUTION TO MINIMIZATION PROBLEM (58)

We now investigate the method of solving problem (58) with $M = 3$, which is an unconstrained optimization. We rewrite the objective function as

$$f(\mathbf{p}_i = \mathbf{p}_o, \mathbf{\Lambda}_i') = \mathbf{p}_i^H \mathbf{\Lambda}_i'^H \mathbf{A}_H \mathbf{A} \mathbf{\Lambda}_i' \mathbf{p}_i = \mathbf{r}^H (\mathbf{\Lambda}_{pi}'\mathbf{A})^H \mathbf{\Lambda}_{pi}' \mathbf{A} \mathbf{r},$$

where $\mathbf{r} = (e^{j\theta_1}, e^{j\theta_2}, e^{j\theta_3})^T$, $\mathbf{\Lambda}_{pi}' = \text{Diag}(\mathbf{p}_o)$. Therefore (58) could also be cast as a quadratic objective optimization with a non-convex constraint

$$\min_{\mathbf{r}} \mathbf{r}^H (\mathbf{\Lambda}_{pi}'\mathbf{A})^H \mathbf{\Lambda}_{pi}' \mathbf{A} \mathbf{r} \ \text{ s.t. } |r_k| = 1, \quad k = 1,2,3.$$

Unfortunately, although the objective function is convex, the feasible set is not. We denote $\mathbf{B} = \mathbf{\Lambda}_{pi}'\mathbf{A} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ and we further rewrite f by extending the quadratic form as

$$f(\mathbf{B},\mathbf{r}) = \sum_{i=1}^{3} \|\mathbf{b}_i\|^2 + \mathbf{b}_1^H\mathbf{b}_2 e^{j(\theta_2-\theta_1)} + \ldots + \mathbf{b}_3^H\mathbf{b}_2 e^{j(\theta_2-\theta_3)}.$$

To minimize f over $\mathbf{r}$ is equivalent to solving

$$\min_{s,t} \ a_1\cos(t+\phi_1) + a_2\cos(s+\phi_2) + a_3\cos(s-t+\phi_3).$$

We could simplify this two-dimensional optimization into a one-dimensional optimization by combining the last two sine function in the following way,

$$a_2\cos(s+\phi_2) + a_3\cos(s-t+\phi_3)$$
$$= \sqrt{[a_2\cos\phi_2 + a_3\cos(\phi_3-t)]^2 + [a_2\sin\phi_2 + a_3\sin(\phi_3-t)]^2}$$
$$\times \cos\left\{ s + \tan^{-1}\left[\frac{a_2\sin\phi_2 + a_3\sin(\phi_3-t)}{a_2\cos\phi_2 + a_3\cos(\phi_3-t)}\right]\right\}. \quad \text{(C.8)}$$

Now a one-dimensional searching over $t \in [0, 2\pi]$ could be implemented to find the optimal solution of (C.8), which in turn could be translated into the optimal solution of (58).

REFERENCES

[1] H. Weingarten, Y. Steinberg, and S. S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[2] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[3] Y. Ma, A. Yamani, N. Yi, and R. Tafazolli, "Low-complexity MU-MIMO nonlinear precoding using degree-2 sparse vector perturbation," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 497–509, Mar. 2016.

[4] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[5] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[6] S. K. Mohammed and E. G. Larsson, "Improving the performance of the zero-forcing multiuser MISO downlink precoder through user grouping," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 811–826, Feb. 2016.

[7] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.

[8] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, pp. 193–228, Nov. 1998.

[9] A. Benjebbour, A. Li, K. Saito, Y. Saito, Y. Kishiyama, and T. Nakamura, "NOMA: From concept to standardization," in *Proc. IEEE CSCN*, Tokyo, Japan, Oct. 2015, pp. 18–23.

[10] H. Lee, S. Kim, and J. Lim, "Multiuser superposition transmission (MUST) for LTE-A systems," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[11] Y. Lan, A. Benjebboiu, X. Chen, A. Li, and H. Jiang, "Considerations on downlink non-orthogonal multiple access (NOMA) combined with closed-loop SU-MIMO," in *Proc. ICSPCS*, Gold Coast, QLD, Australia, Dec. 2014, pp. 1–5.

[12] K. Higuchi and Y. Kishiyama, "Non-orthogonal access with random beamforming and intra-beam SIC for cellular MIMO downlink," in *Proc. IEEE Veh. Technol. Conf.*, Las Vegas, NV, USA, Sep. 2013, pp. 1–5.

[13] J. Kim, J. Koh, J. Kang, K. Lee, and J. Kang, "Design of user clustering and precoding for downlink non-orthogonal multiple access (NOMA)," in *Proc. MILCOM*, Tampa, FL, USA, Oct. 2015, pp. 1170–1175.

[14] C.-H. Liao and H. Morikawa, "Latticing the interference: Non-linear pre-coding for non-orthogonal multiple access in downlink multi-user MIMO system," in *Proc. IEEE PIMRC*, Hong Kong, Aug./Sep. 2015, pp. 710–714.

[15] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[16] Z. Chen, Z. Ding, P. Xu, and X. Dai, "Optimal precoding for a QoS optimization problem in two-user MISO-NOMA downlink," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1263–1266, Jun. 2016.

[17] Z. Chen, Z. Ding, X. Dai, and G. K. Karagiannidis, "On the application of quasi-degradation to MISO-NOMA downlink," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6174–6189, Dec. 2016.

[18] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. Boca Raton, FL, USA: CRC Press, 2001, pp. 18-1–18-33.

[19] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.

[20] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.

[21] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Mar. 2013.

[22] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. Philadelphia, PA, USA: SIAM, 2000.

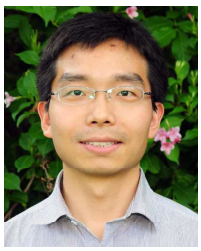
[23] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *Ann. Math. Statist.*, vol. 21, no. 1, pp. 124–127, 1950.
[24] O. Galor, *Discrete Dynamical Systems*. New York, NY, USA: Springer, 2007.
[25] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup beamforming for per-antenna power constrained large-scale arrays," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun./Jul. 2015, pp. 271–275.

Xiaoyan Shi received the B.Sc. degree in electrical and information engineering from Beihang University in 2014. He is currently pursuing the joint Ph.D. degree from the University of Edinburgh and Beihang University. His current research is mainly about beamforming algorithms for MIMO transmission, distributed optimizations for communication networks, as well as some of the applications of these techniques to satellite communications.

John S. Thompson (F'16) is currently a Professor with the School of Engineering, University of Edinburgh. He specializes in antenna array processing, cooperative communications systems, and energy-efficient wireless communications. He has published in excess of 300 papers on these topics. He was a coordinator for the recently completed EU Marie Curie Training Network Advantage, which studies how communications and power engineering can provide future smart grid systems. In 2016, he was elevated to fellow of the IEEE for contributions to antenna arrays and multi-hop communications. He is a Co-Chair of the IEEE SmartGridComm Conference to be held in Aalborg, Denmark. He currently leads two U.K. research projects, which study new concepts for fifth-generation wireless communications. From 2015 to 2017, he has been recognized by Thomson Reuters as a highly cited researcher.

Rongke Liu received the B.S. and Ph.D. degrees from Beihang University in 1996 and 2002, respectively. He was a Visiting Professor with the Florida Institution of Technology, USA, in 2005; The University of Tokyo, Japan, in 2015; and the University of Edinburgh, U.K., in 2018, respectively. He is currently a Full Professor with the School of Electronics and Information Engineering, Beihang University. He received the support of the New Century Excellent Talents Program from the Minister of Education, China. He has attended many special programs, such as China Terrestrial Digital Broadcast Standard. He has published over 100 papers in international conferences and journals. He has been granted 20 patents. His research interest covers wireless multimedia communication, compression coding, channel coding, and aerospace communication.

Majid Safari (S'08–M'11) received the B.Sc. degree in electrical and computer engineering from the University of Tehran, Iran, in 2003, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Iran, in 2005, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2011. He is currently an Assistant Professor with the Institute for Digital Communications, University of Edinburgh. Before joining Edinburgh in 2013, he held the postdoctoral fellowship at McMaster University, Canada. He is currently an Associate Editor of the IEEE COMMUNICATION LETTERS. He was the TPC Co-Chair of the 4th International Workshop on Optical Wireless Communication in 2015. His main research interests include the application of information theory and signal processing in optical communications, including fiber-optic communication, free-space optical communication, visible light communication, and quantum communication.



Pan Cao received the B.Eng. (Hons.) and M.Eng. degrees from Xidian University, China, in 2008 and 2011, respectively, and the Dr.Ing (Ph.D.) degree in electrical engineering from TU Dresden, Germany, in 2015. He was a Post-Doctoral Research Associate with the Institute for Digital Communications, University of Edinburgh, supported by the EPSRC project—Seamless and Efficient Wireless Access for Future Radio Networks—from 2015 to 2017. He was a Visiting Post-Doctoral Researcher at Princeton University, USA, in 2017. He has been a Senior Lecturer in electronics and communications at the University of Hertfordshire, U.K., since 2017. His current research interests include millimeter-wave communication and big data driven optimization for large-scale networks. He received the Best Student Paper Award of the 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications, Cesme, Turkey, in 2012, and the Qualcomm Innovation Fellowship Award in 2013. He currently serves as an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*.